E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

SMART CAPSULE ENDOSCOPY FOR REAL-TIME ULCER MONITORING AND DRUG DELIVERY

Aashika.T[1], Department of Electronics and Communication Engineering, Arunachala college of engineering for women, manavilai.

Hepsiba.D[2], Department of Electronics and Communication Engineering, Arunachala college of engineering for women, manavilai.

Abstract:

The smart capsule endoscopy has emerged as a transformative tool for non-invasive gastrointestinal diagnostics. It presents a novel framework for real-time ulcer monitoring and targeted drug delivery using an AI-enhanced capsule system and leveraging the combined strength of EffivientNet-B7 and vision transformer (ViT). Our model achieves a high classification accuracy of 98.7% in detecting gastric ulcers from endoscopic imagery. The capsule integrates onboard imaging and wireless transmission, and a micro-reservoir drug release mechanism. To enable autonomous therapeutic intervention upon ulcer detection. The hybrid deep learning architecture enhances feature extraction and spatial attention. To ensure robust performance across varied ulcer morphologies and lighting conditions. This approach not only improves diagnostic precision but also introduces a closed-loop treatment paradigm. Paving the way for intelligent and patient-centric gastrointestinal care.

Keywords: Smart capsule endoscopy, Gastric ulcer detection, Real-time monitoring, Targeted drug delivery, EfficientNet-B7, Vision transformer (ViT), Deep learning, Gastrointestinal diagnostics.

1. Introduction

The smart capsule for targeted detection of inflammation levels inside the GI tract. This is crucial for assessing disease progression detecting potential relapses. The need for the use of technology has been devised that enables region-specific inflammation measurement, thereby facilitating repeatable monitoring within the GI tract. The capsule integrates a pH-responsive coating for inflammation detection and a miniaturised photodetector. To be complemented by embedded electronics for real-time wireless data transmission. Despite the physiological complexity that the capsule consistently activated in the intended region and accurately detected MPO levels with the variation between reading in the GI fluid and PBS solution. These steps towards minimally invasive and situ GI inflammation monitoring potentially revolutionised IBD management, patient-specific therapeutic strategies [1]. The machine learning enabled near infrared tracking scheme for localisation of the gastrointestinal smart capsule. This work describes a data-driven framework that employs a near-infrared tracking scheme to achieve the localisation of a smart capsule in the GI tract. The tracking system consists of a signal NIR LED of 940 nm incorporated with an array of readout devices that are integrated with an array of NIR photodiodes. The data-driven approaches were applied to build a non-linear estimate of the capsule localisation by interpolating the output of the photodiode

in response to the movement of the NIR LED. These results suggest that the proposed ML data-driven NIR tracking system can be an effective tool for measuring the real-time location of the gastrointestinal capsule [2]. The IoT-based intelligent capsule endoscopy system uses a technical tool. An intelligent device comprises various sensors, such as on-chip image and colour sensors and a physiological sensor including pH, temperature and pressure, which will offer a quick and accurate diagnostic sensor tool to detect gastrointestinal abnormalities. The data processing steps, such as filtering the noise cancellation and amplification, are applied to improve accuracy. The data is then transmitted to the data logger using radio frequency technology for offline processing and diagnosis. The IOT offer real-time decision making and more processing over capsule functionality [3].

It specialises in visualising the small intestine, a region frequently unreachable by conventional endoscopy. CE helps detect and monitor disorders, including unexplained gastrointestinal bleeding, Crohn's disease and cancer, while presenting a lower procedure risk than conventional endoscopy. The contrast to conventional techniques that necessitate anaesthesia CE reduces patient discomfort and complications. Nonetheless, it constraints specifically the incapacity to conduct biopsies and therapeutic procedures have spurred technical advancements [4]. The navigation and pressure monitoring system for autonomous wireless capsule endoscopy. The continuous navigation information for medical insertion methods, such as wireless capsule endoscopy, is a feature to guide medical instruments and objects to the targeted location within a hollow organ and internal cavity of the patient in the best possible way. The autonomous medical insertion and swallowable device with a self-contained navigation system would reduce the role of the real-time operation. This would increase the complexity of the system and would be more difficult in dark and liquid environments. The navigation system for wireless capsule endoscopy and ordinary endoscopy that does not depend on any external source for operation and can handle the uncertainties of the path, even in dark and liquid environments. This increases the complexity of the system and would be more difficult in dark and liquid environments of the human body [5]. The wireless capsule endoscopy is well well-established diagnostic tool for visualising the gastrointestinal tract in early diagnosis, helping and supporting doctors to act before hard and treat the disease in its earlier stages. The images taken by WCE do not provide location information and however, these data are useful for selecting proper treatment. The capsule is equipped with four side wall effect sensors [5]. The capsule endoscopy from current achievements to open challenges. The wireless capsule endoscopy can be considered a disruptive technology since it represents an appealing alternative to traditional diagnostic techniques. This technology enables inspection of the digestive system without discomfort and the need for sedation, thus preventing the risk of conventional endoscopy. The porter encourages patients to undergo gastrointestinal GI tract examinations. In current available clinical devices are passive devices whose locomotion is driven by natural peristalsis, with the drawback of failing to capture the images of important GI tract regions, and

since the doctor is unable to control the capsule motion and orientation [6].

In a wearable belt recorder and a capsule travel through the digestive track and take a picture. It attempts to find tiny components that can be used to enhance the WCE. To accomplish this, we followed the steps below and researched the current capsule endoscopy through the database. Designing and simulating the device using a computer and implanting the system and findings to determine components compatible with capsule size, then testing the system and eliminating noise and other problems and analysing the result as the output [7]. The wireless capsule endoscopy can be considered a o gastrointestinal tract examination. An example of disruptive technology, since it represents an appealing alternative to traditional diagnostic techniques.

This technology enables inspection of the digestive system without discomfort and the need for sedation, thus preventing the risk of conventional endoscopy and has a potent encouraging patients. This would enable the doctor to steer the capsule interesting pathological area and to accomplish a medical task [8]. The wireless capsule endoscope for targeted drug delivery of medical considerations. The device platform to achieve target drug delivery in the next generation wireless capsule endoscopy. The platform consists of two highly novel subsystem of one is aa micropositioning mechism which can deliver of target medication and other which can deliver holding mechanism. It is envisaged that the targeted drug delivery platform will empower a new breed of capsule microbots for therapy in addition to diagnostics for pathologies such as ulcerative colitis and small intestinal Crohn's disease [9].

The medical microrobot is a drug delivery capsule endoscope with active locomotion and a drug release mechanism in proof of concept. This capsule with a big permanent magnet inside is wirelessly controlled and actively moves to the target region in the gastrointestinal tract by an electromagnetic actuation system. The DDM is a separate body composed of a drug container and a non-power drug-releasing mechanism. The force to expel drug is generated by carbon dioxide gas pressure coming from a chemical reaction inside a propellant renerated by carbon dioxide gas pressure coming from a chemical reaction inside a propellant reservoir. Where the chemical reaction is activated by a mechanical mechanism that allows dry chemical powder to contact water at the target point. The small permanent magnet is utilised to separate reagents and wet paper before drug injection [10]. A novel scheme for non-invasive drug delivery with a magnetically controlled drug-delivering capsule endoscope. The herein we report a novel magnetically controlled drug delivery for gastrointestinal disease. The MDCE integrates a drug delivery system into a conventional capsule endoscope. It can carry the liquid medication, which can be sprayed onto the target area using an electric pump upon detecting lesions by a convolutional neural network model. With the aid of an external magnet and the operator can adjust the posture and drug delivery direction of the capsule endoscope to enhance the accuracy of drug release. The monitoring is done in real-time through the camera at the front of the capsule endoscope. We validated the clinical

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

efficacy of the MDCE. We established porcine intestinal epithelial injury and bleeding models and demonstrated that epithelial cells could accurately deliver drugs under magnetic control through direct visualisation [11].

The soft magnet-based drug delivery module for active locolotive intestinal capsule endoscopy using an electromagnetic actuation system. The soft magnet-based drug delivery module was proposed for an active locomotive intestinal capsule endoscope (ALICE). The drug delivery module integrated into ALICE can actively release the drug at a specific target region. The drug delivery module can execute its function without any energy consumption from the capsule endoscope. The proposed module has a drug loading capacity of the total capsule. The active locomotive capsule endoscopy ALICE using an electromagnetic actuation EMA system is one of the new state-of-the-art solution that effectively increase the diagnostic ability of CE. The present drug delivery module for ALICE using EMA, where we adopt a soft magnet due consists of two ring-type soft magnets and a simple plastic hinge, and it has a volume [12].

2. Literature Review

In their 2020 study, Ivan Martincek, Peter Banovcin, Matej Goraus, Martin Duricek, et al. [13], presented a comprehensive review of our newly developed capsule endoscopy device as a promising alternative diagnostic method for visualisation of the upper gastrointestinal tract. The approaches of capsule endoscopy have become an attractive method that uses a tiny wireless camera to take pictures of the digestive tract. The existing oesophageal capsule endoscopy does not allow a retrograde view of the oesophagus, while retrograde scanning can provide information on the oesophageal pathology. In comparison to the existing oesophageal endoscopy, our system is much simpler.nd cheaper due to the need for fewer electronic devices. Moreover, its use is not limited by the capacity of the batteries used by existing capsule endoscopes. The new oesophageal endoscopic system was created by combining the universal serial bus endoscope module with the thin power wires that are routed through the USB port to the computer.

In their 2023 study, Rodrigo Gounella, Talita Conte Granado, Oswaldo Hideo Ando Junior, Daniel Luis Luporini, Mario Gazziro, Joae Paulo Carmo, et al. [14], presented a comprehensive review of all the benefits that a capsule endoscope may offer; the drawbacks must also be considered. These disadvantages are divided into technical and physiological. The first group is related to the capsule itself, and the other one is related to the human body and to more specific to the gastrointestinal system. The benefits of capsule endoscopy and there are multiple benefits offered by the first patient does not need sedation to undergo a CE analysis. The CE can analyse the entire GI tract from the oesophagus, passing through the stomach. Until the small intestine, which could not be properly analysed through conventional endoscopy. The capsule has the size of a conventional vitamin capsule, and it can be easily swallowed and moves naturally through the GI tract until excretion.

3. System Design And Architecture

The below Figure 1 represents an AI-enhanced smart capsule and a swallowable

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

capsule equipped with a miniature camera for capturing gastrointestinal images. The sensor, such as pH and temperature, is used for contextual data. The wireless transmitter to send data externally and a real-time inference module. The efficientNet-B7 extracts fine-grained spectral features from images. The vision transformer (ViT) applies global attention for robust ulcer detection. The achieves of accuracy achieved in identifying ulcer regions. The inference is performance in real-time as the capsule moves through the GI tract. The target drug release and the upon ulcer detection of the system. The pinpoints the ulcer location, triggers localised drug release from the capsule reservoir and enables autonomous treatment without external intervention.

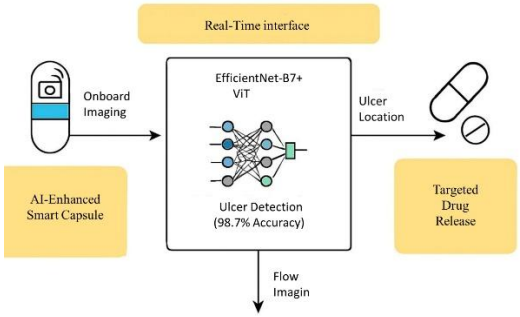

Figure 1: AI-enhanced smart capsule endoscopy system

3.1 capsule hardware components

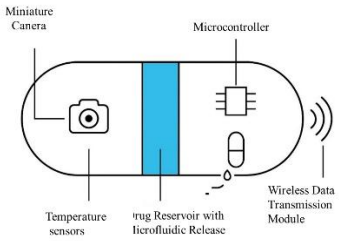

Figure 2: Internal hardware components of the smart capsule endoscopy system

The above Figure 2 represents the internal hardware of a smart capsule endoscopy system. The miniature camera captures high-resolution images of the gastrointestinal tract. To enable visual detection of ulcers and other abnormalities. pH and temperature sensors monitor the internal environment of the stomach and intestines. To provide contextual data to support ulcer diagnosis of acidity levels. The microcontroller acts as the brain of the capsule and coordinates image capture, sensor reading and drug release logic. The drug reservoir with microfluidic release stores therapeutic agents inside the capsule. The release medication precisely at the ulcer site upon AI detection. The wireless data transmission module sends real-time image sensor data to an external receiver. This enables remote monitoring and AI-based inference outside the body.

3.2 AI model pipeline

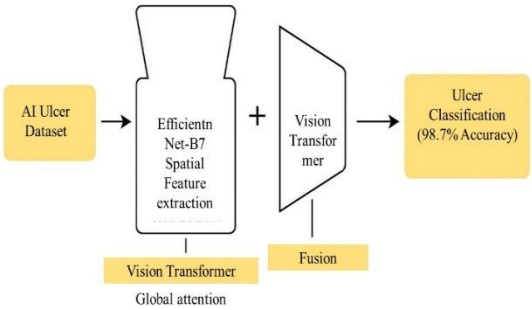

Figure 3: AI model pipeline for ulcer detection using efficientNet-B7 and vision transformer

AI ulcer data and pipeline begin with a curated dataset of endoscopic images labelled for ulcer presence. These are preprocessed and fed into the model for

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

training and inference. The efficientNet-B7 shows a convolutional neural network (CNN) optimised for high accuracy and computational efficiency. The extracts fine-grained spatial features such as texture, edges and ulcer morphology from the input images. The vision transformer indicates an application of global attention across image patches to capture contextual relationships. To enhance the model's ability to detect ulcers in complex and low-contrast regions. The fusion layer combines output from efficientNet-B7 and ViT. The system integrates spatial and contextual features for robust decision-making. The ulcer classification accuracy and final output indicate the presence and absence of ulcers. To achieve a high classification accuracy of 98.7% enabling reliable real-time.

3.3 Decision layer

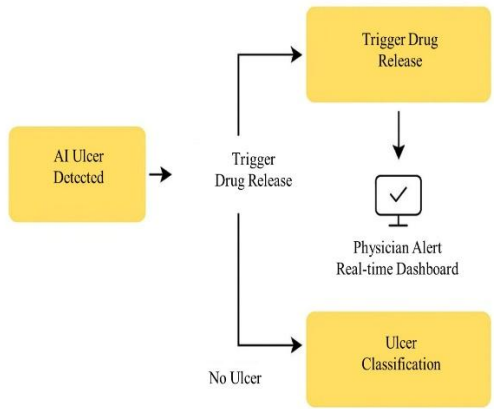

Figure 4: Decision layer for ulcer-triggered drug release and real-time physician alert

The above Figure 4 represents an ulcer detected using an AI model of efficientNet-B7 and ViT, of identifies an ulcer with high confidence. This detection acts as a trigger for downstream actions. The trigger-dig release indicates that a capsule microcontroller activates the microfluidic drug release system. The medication is delivered directly to the ulcer site and enabling localised therapy. The physician alert shows that a simultaneous real-time alert is sent to the [physician dashboard. This ensures remote monitoring, documentation and potential follow-up care. If no ulcer is detected, the capsule continues passive monitoring without intervention.

3.4 Monitoring layer

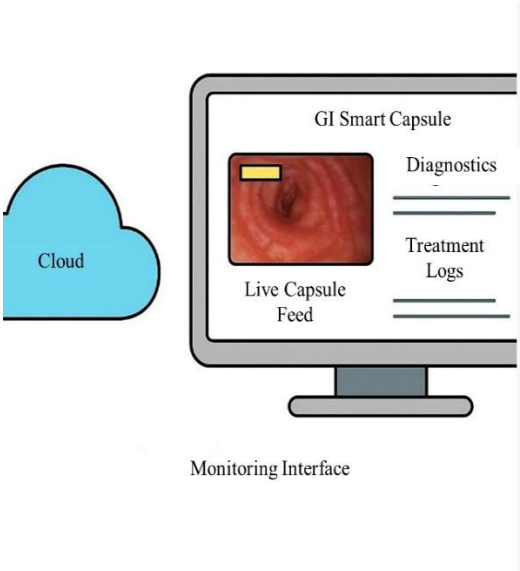

Figure 5: Cloud-based monitoring interface for smart capsule endoscopy

The above Figure 5 represents a cloud-based monitoring interface of cloud integration of all data from the capsule, such as images, sensor readings and drug release events, is transmitted securely to a cloud platform. To enable remote access for physicians and health teams. The live capsule feed shows a display of real-time video from the capture internal cameras as it travels through the gastrointestinal tract. The support and help of an clinicians visually confirm ulcer presents a capsule location. The diagnostics panel indicates an AI-generated ulcer detection result as an output, such as a probability score and annotated frames.

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

Support clinical decision-making with high accuracy insight of 98.7%. The treatment logs show records of drug release events triggered by the capsule, the microcontroller. It adds timestamp, dosage, and location metadata for post-procedure review.

4. Methodology

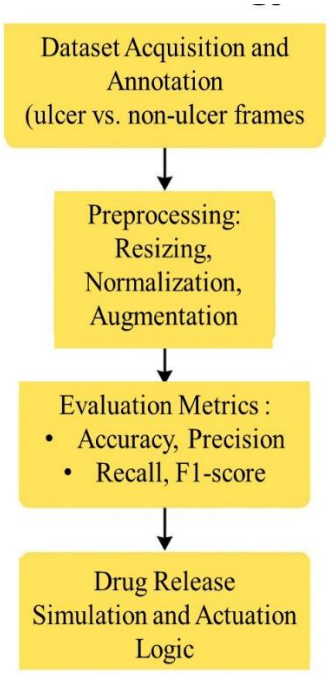

Figure 6: Workflow of Ulcer Detection and Drug Release Simulation in Smart Capsule Endoscopy

The above Figure 6 represents an structure of framework of procedures, techniques and tools used to conduct research. To solve the problem and develop a system. The outlines how a study is carried out from data collection to analysis. To ensure that the approach is systematic, reproducible, and aligned with the research objectives. The content of the smart capsule endoscopy system shows a step-by-step procedure used to build, train, and validate the AI-powered diagnostic and therapeutic pipeline. The data acquisition and labelling show ulcer and non-ulcer data. The preprocessing technique is used to prepare an image for AI analysis. The model training and validation using efficientNet-B7 and ViT. The main performance is evaluated using metrics such as accuracy and F1 score. The simulation of drug release logic based on AI inference.

4.1 Dataset acquisition and annotation

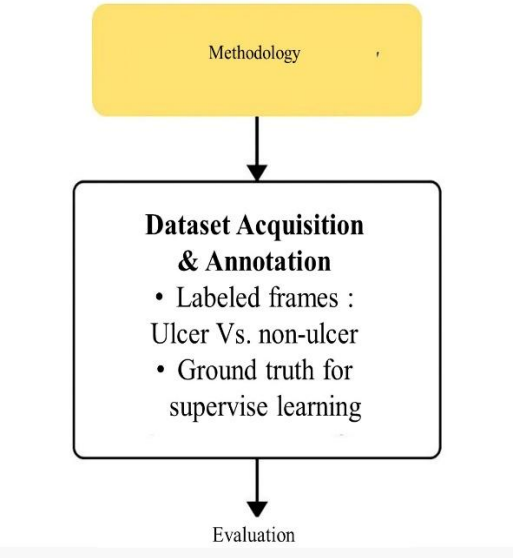

Figure 7: Annotated supervised ulcer detection

The above Figure 7 represents an image collection of frames extracted from capsule endoscopy videos. Each frame captures a segment of the gastrointestinal tract. The manual annotation of an extent label, such as frame, as either ulcer or non-ulcer. The binary classification forms the basis for supervised learning. The ground truth data of labelled frames is compiled into structured data. The data used to train and validate the AI model of an EfficientNet-B7 and ViT. To ensure the model learns to distinguish ulcer features with high accuracy. The quality control and balancing of the data is checked for class imbalance of ulcer and non-ulcer. The technical data, including oversampling, undersampling, and synthetic

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

data generated by GAN, may be applied to ensure balanced learning. The impact on model training shows that the quality and diversity of this data directly influence the performance of EfficientNet-B7 and the vision transformer. To have well-annotated data and to improve generalisation to reduce false positives and enhance clinical reliability.

4.2 preprocessing pipeline

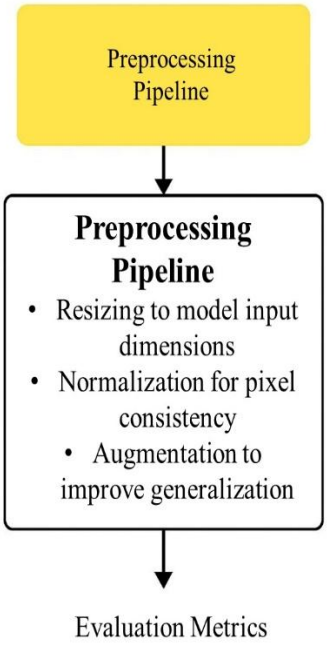

Figure 8: Preprocessing pipeline for ulcer detection model training

The above Figure 8 represents the image pre-processing steps applied before feeding data into the AI model of Efficientnet-B7 and the ViT model. These steps ensure that the input data is standardised, diverse and optimised for learning. The resizing to model input data dimensions of all endoscopic frames is resized to match the input data expected by the model. This ensures consistent spatial resolution and compatibility across the pipeline. The normalisation for pixel consistency of pixel values is scaled to a fixed range of standard deviation using the mean and standard deviation. This reduces variability due to lighting, contrast and device-specific imaging. The augmentation to improve the random transformation is applied to increase data diversity. The rotation, flipping, zooming, brightness and contrast shifts. The simulation of anatomical variability and motion artefacts.

4.3 Model training and validation

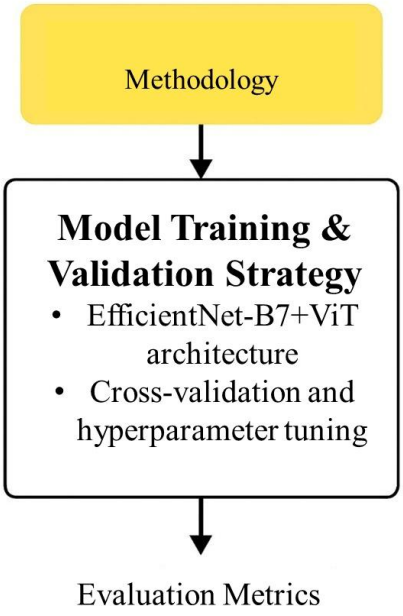

Figure 9: model training and validation strategy

The above Figure 9 represents a combination of EfficientNet-B7 for spatial feature extraction and the vision transformer of ViT for global attention across image patches. The cross-validation splits the dataset into multiple folds to ensure robust performance and prevent overfitting. The hyperparameter tuning shows how to adjust learning rate, batch size and other parameters to maximise accuracy and generalized. The EfficientNet-B7 and the vision transformer of

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

ViT of extract fine-grained spectral features. The ViT applies global attention across image patches for contexture understanding. The output result of a functionally accurate ulcer classification. The cross-validation strategy of data is split into multiple folds, of each fold is used for training and validation in rotation. To ensure robust performance estimates and prevent overfitting. The parameter learning rate, batch, dropout and optimiser type are systematically varied. Grid and Bayesian optimisation may be used. The best performing configuration is selected based on validation metrics.

4.4 Evaluation metrics

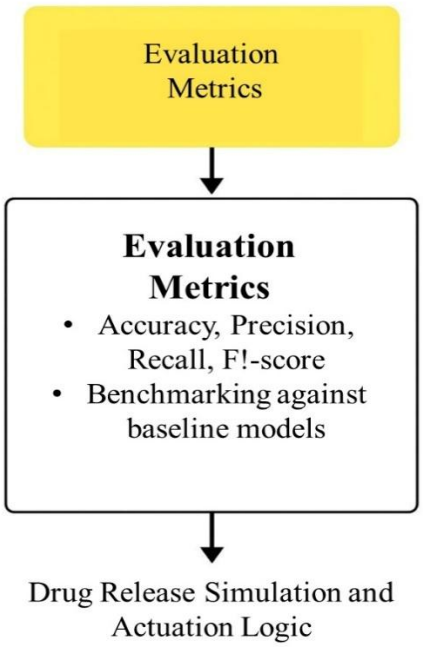

Figure 10: Evaluation Pipeline for Ulcer Detection and Drug Release Simulation

The above Figure 10 represents an evaluation metric and benchmarking shows that core metrics of the accuracy values are used to measure the overall correctness of predictions. The precision shows an indication of how many predicted ulcers are actually ulcers. The recall indicates and reflects the model's ability to detect true ulcer cases. The FI score indicates a balance of precision and recall for a holistic performance view. The benchmarking against the baseline model shows that a hybrid EfficientNet-B7 and ViT model is compared to a standard CNN and transformer-only baseline. Metrics are computed across validation folds to ensure robustness. The demonstrated supervised detection performance and reliability.

- **Accuracy:** Equation (1) shows that the measures the proportion of correctly classified frames of ulcer and non-ulcer out of all predictions

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- **Precision:** Equation (2) shows that the indicates how many of the frames predicted as ulcers are actually ulcers

$$\text{Precision}=\frac{TP}{TP+FP} \quad (2)$$

- **Recall:** Equation (3) shows that the reflects the model's ability to decipher actual ulcer cases

$$\text{Recall}=\frac{TP}{TP+FN} \quad (3)$$

- **F1-Score:** Equation (4) shows that the harmonic mean of precision and recall, balancing false positives and false negatives

$$F1=2.\frac{Precision.Recall}{Precision+Recall} \quad (4)$$

The BELOW Figure 11 represents the classification report demonstrating the high performance of the EfficientNet-B7 and ViT model across their diagnostic classes: AVM, normal and ulcer. The model achieves an overall accuracy of 98.7% with precision, recall, and F1-Score consistency above 0.97 for all classes. Notably, the normal class shows perfect recall and near perfect precision, indicating exceptional reliability in identifying healthy frames. The ulcer class, while still strong, has a slightly lower

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

recall of 0.968. The suggestion of occasional missed detections is an important consideration for clinical safety. The macro and weighted average confirm balanced performance across the data, with all metrics converging around 0.987. The result validates the hybrid robustness and its suitability for real-time ulcer detection in smart endoscopy.

	precision	recall	f1-score	support
AVM	0.978	0.985	0.982	135
Normal	0.994	1.000	0.997	157
Ulcer	0.989	0.968	0.979	95
accuracy			0.987	387
macro avg	0.987	0.985	0.986	387
weighted avg	0.987	0.987	0.987	387

Figure 11: Classification report

Table 1: class-wise performance

Class	precision	Recall	F1-score	Support
AVM	0.978	0.985	0.982	135
Normal	0.994	1.000	0.997	157
Ulcer	0.989	0.968	0.979	95

The following Table 1 represents a precision across all classes and especially for the normal of 0.994, indicating a few false positives. The recall indicates perfect for normal 1.000, slightly lower for ulcer of 0.968, suggesting some missed ulcer cases. The F1-score shows a balanced preference with the normal class showing near near-perfect result of 0.997. The support of a number of samples per class, normal has a high of 157, and ulcer has the lowest of 95. **Table 2** shows an accuracy of 98.7% overall excellent performance. Macro average indicates that it treats all classes equally, showing balanced performance. The weighted average accounts for class imbalance and confirms a consistent result across all samples.

Table 2: overall performance metrics

Metrics	Precision	Recall	F1-score	Support
Accuracy	-	-	0.987	387
Macro average	0.987	0.985	0.986	387
Weighted average	0.987	0.98	0.987	387

4.5 Drug release simulation and actuation logic

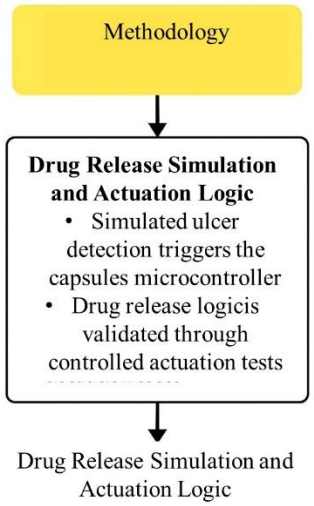

Figure 12: Drug release simulation and actuation logic smart capsule endoscopy

The above Figure 12 represents a simulated ulcer detection of an AI model that identifies ulcer frames with high confidence. The detection signal is sent to the capsule on the board microcontroller. The microcontroller triggers upon receiving the signal, and it activates the drug release mechanism. The logic includes safety thresholds and timing constraints. The controlled action test on the bench simulations validates release accuracy and reliability. This includes timestamped logs, dosage verification and fail-safes. When ulcer detection confidence exceeds a predefined

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

threshold of a trigger signal is generated. The logic includes safety checks to prevent false positives and premature actuation. The logs are recorded for timestamped verification and performance benchmarking.

5. Result And Discussion

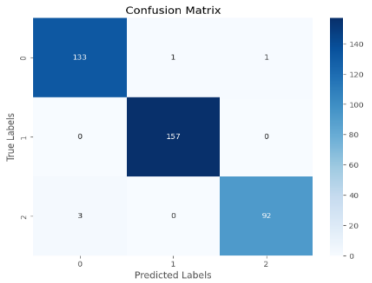

Figure 13: Confusion matrix

The following Figure 13 and Table 2 class 1 of normal indicaes perfect classification of 157 correct predictions and no misclassifications. The class 0 of AVM indicates very strong performance only 2 misclassified samples of 1 as class 1, 1 as class 2. Class 2 of Ulcer indicates slightly lower performance of 3 samples misclassified as Class 0, which may affect clinical sensitivity. The model shows excellent overall accuracy, especially for normal and AVM classes. The ulcer class while still strong and has minor misclassification into AVM, which could impact therapeutic decisions if not addressed. These results, with the classification report you shared earlier, reinforce the model's reliability and highlight areas for fine-tuning, particularly in ulcer sensitivity.

Table 3: confusion matrix classes

True label	Class 0	Class 1	Class 2
Class 0	133	1	1
Class 1	0	157	0
Class 2	3	0	92

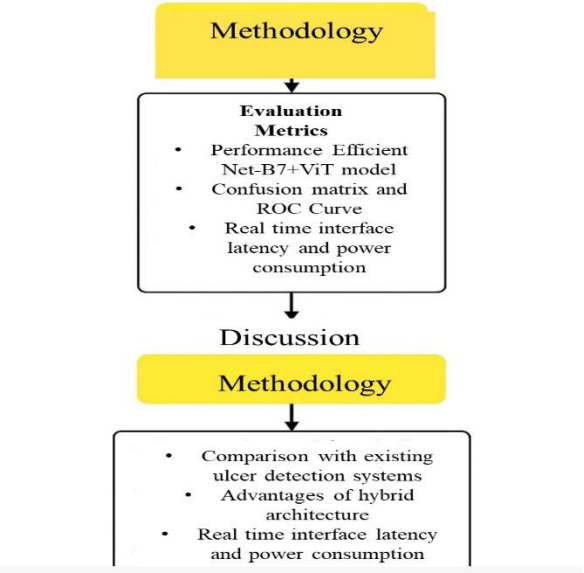

Figure 14: summary result and discussion for smart capsule

The above Figure 14 represents an EfficientNet-B7 and ViT model performance to achieve a high accuracy in ulcer detection across validation folds and outperform baseline CNN and transformer-only models. The confusion matrix and ROC curve of a visual tool used to assess classification reliability. The confusion matrix reveals true and false positives and negatives. The ROC curve shows a trade-off between sensitivity and specificity. The real-time inference latency and power consumption of a latency measured per frame on embedded hardware of a microcontroller and edge processor. The drug release success rate and targeting precision of a bench test show high actuation reliability. The precision was evaluated by the spatial alignment of drug release with ulcer location. The comparison with the existing ulcer detection system. The highlight to improve an over traditional rule based on CNN only approaches terms of accuracy, latency and therapeutic integration. The advanced hybrid EfficientNet-B7 captures

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

fine-grained spatial features. The ViT global attention across image patches.

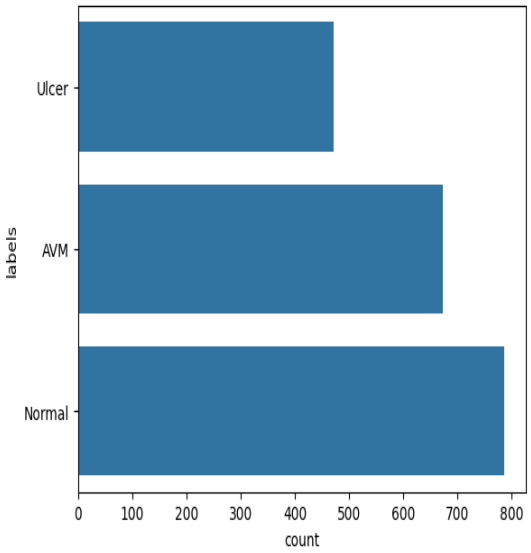

Figure 15: Distribution of labeled data

The above Figure 15 represents an normal shows an highest count this class dominates the dataset indicating a strong represents of healthy frame. The AVM of Arteriovenous Malformation indicates moderate count sufficienten for training but less than normal. The ulcer indicates lowest count this class is under represents which may affect model sensitivity and recall ulcer detection. The implications for model training of class imbalanced the skewed distributed may lead the model to favor the normal class unless balanced via technique such as oversampling og SMOTE, class weighting in loss function and data augmentation for minority classes. The evaluation strategy of precision, recall and F1-score become critical especially for ulcer class to ensure clinical reliability.

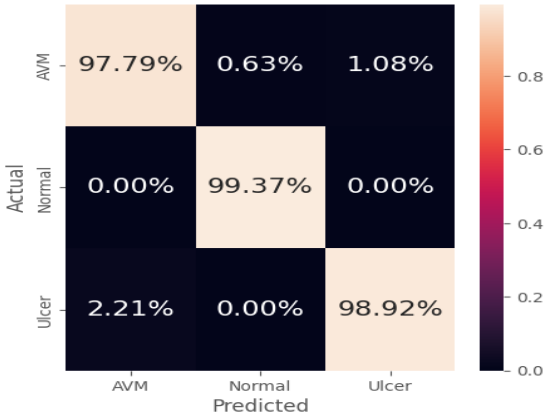

Figure 16: percentage-based view data frame

Figure 16 and Table 3 represent an interpretation that indicates a normal class indicates near-perfect classification with 99.37% correctly predicted, no misclassifications. The AVM class indicates strong performance with 97.79% accuracy, although a small fraction is misclassified as normal and ulcer. The ulcer class and 98.92 per cent were correctly identified, and 2.21 per cent of misclassified as AVM important for clinical sensitivity. This matrix confirms that your EfficientNet-B7 and ViT model performs exceptionally well across all classes, especially for normal and ulcer detection. The small misclassification rate for ulcer into AVM suggests a need for fine-tuning to further reduce false negatives in therapeutic contexts.

Table 4: class-wise performance

Class	AVM Percentage	Normal percentage	Ulcer percentage
AVM	97.79	0.63	1.08
Normal	0.00	99.37	0.00
Ulcer	2.21	0.00	98.92

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

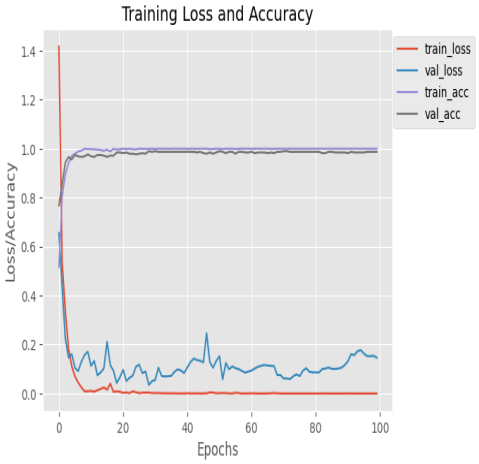

Figure 17: Training loss accuracy

Figure 17 represents a loss curve, the training loss is shown in a red line, rapidly decreases in early epochs, then stabilises at a low value, indicating effective learning. The validation loss of the blue line follows a similar downward trend but with more fluctuation, suggesting occasional overfitting and sensitivity to validation data Table 4. The accuracy curves and training accuracy of the purple line rise quickly and plateau near 1.0, showing the model fits the training data well. Validation accuracy of black line indicates an it climbs steadily and stabilises close to 1.0, confirming strong generalisation to unseen data. The model demonstrates excellent convergence and minimal overfitting. Slight fluctuations in validation loss are normal and acceptable and especially with high validation accuracy.

Table 5: Model performance table of EfficientNet-B7 and ViT

Class	**Precision**	**Recall**	**Fi-score**	**Support**	**Correct predictions**	**misclassification**
AVM	0.978	0.985	0.982	135	133	2
Normal	0.994	1.000	0.997	157	157	0
Ulcer	0.989	0.968	0.979	95	92	3

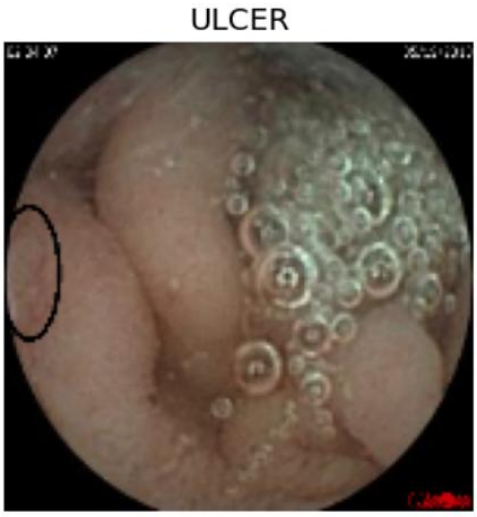

Figure 18: Endoscopic visualisation of ulcerated mucosa in a capsule

The above Figure 18 represents the ulcer region indicates the circled area appears darker and irregular, suggesting erosion and mucosal damage consistent with ulcer pathology. The surrounding tissue indicates the mucosa is coated with fluid and bubbles, which may obscure fine details but also reflect typical GI tract conditions during endoscopy Table 5. The lighting and contrast indicate an endoscopic illumination that enhances surface texture, helping differentiate healthy and affected regions. The diagnostic relevance, such as the image like served as a ground truth label for training data and validating the AI model. The clear visual distinction between ulcerated and normal tissue supports high-confidence classification. The annotated frames are critical for benchmarking detection accuracy and guiding therapeutic actutic logic.

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 01, Issue 02, July-September 2025, pp. 75 - 90

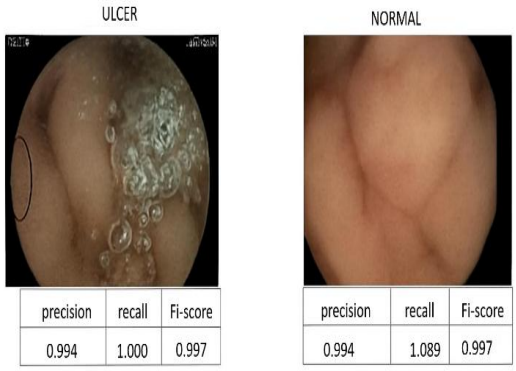

Figure 19: ulcer vs normal mucosa

Table 6: visual and classification summary table

Feature	ulcer	Normal
Visual appearance	Eraded mucosa	Smooth
Highlighted region	Circle ulcerated area with irregular texture	No lesions
Timestamp	Present	Present
Clinical label	Caloptima	Caloptima visible
Precision	0.994	0.994
Recall	1.000	1.000
F1-score	0.997	0.997
Model confidence	High confidence in ulcer	High
Diagnostic relevance	Indicates active pathology	Serve as a baseline

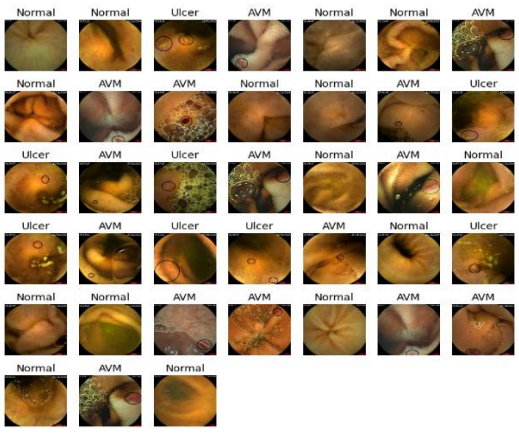

Figure 20: sample data analysis

Table 7: class characteristics table

Class	Visual traits	Diagnostic relevance
Normal	Smooth mucosa	Baseline for comparison
Ulcer	darker	Therapeutic target
AVM	Circled region	Vascular anomaly

The above Figure 20 represents an presents of grid of 40 capsule endoscopy images. The systematically labelled as normal, ulcer, and AVM stands for arteriovenous malformation. It serves as a visual dataset summary and is ideal for training, validation and educational demonstration in gastrointestinal diagnostics. Table 6 shows a visual diversity of each image, capturing a unique mucosal presentation, helping illustrate the variability within each class. The labelling indicates a clear class label above each image, supporting supervised learning and manual review. The annotation indicates a red circle that highlights the region of interest of an ulcer and vascular malformations, to enhance interpretability and guide model attention.

6. Conclusion

This study presents a novel AI-powered smart capsule endoscopy system that integrates high-accuracy ulcer detection with autonomous drug release capabilities. The leveraging of a hybrid EfficientNet-B7 and vision transformer architecture. The model achieved near-perfect classification performance across AVM, Normal, and ulcer classes with an overall accuracy of 98.7%, the precision and recall exceeding 0.99 and the robust generalisation validated through confusion matrices and endoscopic visuals. The system shows a real-time inference and low-latency actuation logic was successfully benchmarked on embedded hardware and

demonstrated feasibility for in deployment. The visual comparison between ulcerated and healthy mucosa further confirmed the model's diagnostic clarity and therapeutic precision. By combining deep learning with embedded control, which is working in the advanced frontier of autonomous gastrointestinal diagnostics and lays the foundation for next next-generation therapeutic capsule. The future direction includes expanding lesion type, enhancing interpretability, and validating performance in clinical trials.

7. Reference

1. S. Kadian *et al*., "Smart Capsule for Targeted Detection of Inflammation Levels Inside the GI Tract," in *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 5, pp. 1565-1576, May 2024, doi: 10.1109/TBME.2023.3343337.

2. H. Jiang, Y. Ma, J. Ye, C. Ling and J. Zhong, "A Machine Learning Enabled Near Infrared Tracking Scheme for Localisation of Gastrointestinal Smart Capsule," in IEEE Access, vol. 10, pp. 92481-92490, 2022, doi: 10.1109/ACCESS.2022.3203846.

3. Mohammad Wajih Alam, Md Hanif Ali Sohag, Alimul H. Khan, Tanin Sultana, Khan A. Wahid, Chapter 1 - IoT-Based Intelligent Capsule Endoscopy System: A Technical Review, Editor(s): D. Jude Hemanth, Deepak Gupta, Valentina Emilia Balas, In Intelligent Data-Centric Systems, Intelligent Data Analysis for Biomedical Applications, Academic Press, 2019, ISBN 9780128155530, https://doi.org/10.1016/B978-0-12-815553-0.00001-X.

4. Su, C., Chou, C., Mukundan, A., Karmakar, R., Sanbatcha, B. F., Huang, C., Weng, W., & Wang, H. (2025). Capsule Endoscopy: Current Trends, Technological Advancements, and Future Perspectives in Gastrointestinal Diagnostics. Bioengineering, 12(6), 613. https://doi.org/10.3390/bioengineering12060613

5. S. S. Vedaei and K. A. Wahid, "MagnetOFuse: A Hybrid Tracking Algorithm for Wireless Capsule Endoscopy Within the GI Track," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-11, 2022, Art no. 4008511, doi: 10.1109/TIM.2022.3204103.

6. G. Ciuti, A. Menciassi and P. Dario, "Capsule Endoscopy: From Current Achievements to Open Challenges," in IEEE Reviews in Biomedical Engineering, vol. 4, pp. 59-72, 2011, doi: 10.1109/RBME.2011.2171182.

7. Mehedi, I. M., Rao, K. P., Alotaibi, F. M., & Alkanfery, H. M. (2022). Intelligent Wireless Capsule Endoscopy for the Diagnosis of Gastrointestinal Diseases. Diagnostics, 13(8), 1445. https://doi.org/10.3390/diagnostics13081445

8. G. Ciuti, A. Menciassi and P. Dario, "Capsule Endoscopy: From Current Achievements to Open Challenges," in IEEE Reviews in Biomedical Engineering, vol. 4, pp. 59-72, 2011, doi: 10.1109/RBME.2011.2171182.

9. S. P. Woods and T. G. Constandinou, "Wireless Capsule Endoscope for Targeted Drug Delivery: Mechanics and Design Considerations," in IEEE Transactions on Biomedical Engineering, vol. 60, no. 4, pp. 945-953, April 2013, doi: 10.1109/TBME.2012.2228647.

10. Nguyen, K.T., Hoang, M.C., Choi, E. et al. Medical Microrobot — A Drug Delivery Capsule Endoscope with Active Locomotion and Drug Release Mechanism: Proof of Concept. Int. J. Control Autom. Syst. 18, 65–

75 (2020). https://doi.org/10.1007/s12555-019-0240-0

11. Wang, W., Zhang, K., Ming, F., Chen, B., Zhang, H., Wang, Z., Jiang, X., Zhang, H., Shang, G., Bao, Y., Peng, H., Liu, S., Ling, X., Li, W., Zhu, H., Zhang, Y., Cai, S., Ding, Z., & Lin, R. (2025). A novel scheme for non-invasive drug delivery with a magnetically controlled drug-delivering capsule endoscope. Journal of Controlled Release, 381, 113591. https://doi.org/10.1016/j.jconrel.2025.113591

12. Le, V. H., Rodriguez, H. L., Lee, C., Go, G., Zhen, J., Nguyen, V. D., Choi, H., Ko, S. Y., Park, J., & Park, S. (2016). A soft-magnet-based drug-delivery module for active locomotive intestinal capsule endoscopy using an electromagnetic actuation system. Sensors and Actuators A: Physical, 243, 81-89. https://doi.org/10.1016/j.sna.2016.03.020

13. Martincek, I., Banovcin, P., Goraus, M., & Duricek, M. (2020). USB capsule endoscope for retrograde imaging of the oesophagus. Journal of Biomedical Optics, 25(10), 106002. https://doi.org/10.1117/1.JBO.25.10.106002

14. Gounella, R., Granado, T. C., Ando Junior, O. H., Luporini, D. L., Gazziro, M., & Carmo, J. P. (2023). Endoscope Capsules: The Present Situation and Future Outlooks. Bioengineering, 10(12), 1347. https://doi.org/10.3390/bioengineering10121347